What is claimed is:

1. A method for music analysis comprising the steps of:

acquiring a music soundtrack;

re-sampling an audio stream of the music soundtrack so that the re-sampled audio stream is composed of blocks;

applying Fourier Transformation to each of the blocks;

deriving a first vector from each of the transformed blocks, wherein components of the first vector are energy summations of the block within a plurality of first sub-bands;

applying auto-correlation to each sequence composed of the components of the first vectors of all the blocks in the same first sub-band using a plurality of tempo values, wherein, for each sequence, a largest correlation result is identified as a confidence value and the tempo value generating the largest correlation result is identified as an estimated tempo; and

comparing the confidence values of all the sequences to identify the estimated tempo corresponding to the largest confidence value as a final estimated tempo.

2. The method as claimed in claim 1 further comprising the step of:

deriving a second vector from each of the transformed blocks, wherein components of the second vector

are energy summations of the block within a plurality of second sub-bands; and

detecting micro-changes using the second vectors.

3. The method as claimed in claim 2, wherein, for each block, a micro-change value which is a sum of differences between the second vectors of the block and previous blocks is calculated.

4. The method as claimed in claim 3, wherein each micro-change value is derived by the following equation:

$$MV_{(n)} = Sum(Diff(V2_{(n)},V2_{(n-1)}),Diff(V2_{(n)},V2_{(n-2)}),Diff(V2_{(n)},V2_{(n-3)}),Diff(V2_{(n)},V2_{(n-4)})),$$

where $MV_{(n)}$ is the micro-change value of the n^{th} block, $V2_{(n)}$ is the second vector of the n^{th} block, $V2_{(n-1)}$ is the second vector of the $(n-1)^{th}$ block, $V2_{(n-2)}$ is the second vector of the $(n-2)^{th}$ block, $V2_{(n-3)}$ is the second vector of the $(n-3)^{th}$ block and $V2_{(n-4)}$ is the second vector of the $(n-4)^{th}$ block.

5. The method as claimed in claim 4, wherein the difference between two of the second vectors is a difference of amplitudes thereof.

6. The method as claimed in claim 5, wherein the micro-change values are compared to a predetermined threshold, and the blocks having the micro-change values larger than the threshold are identified as micro-changes.

7. The method as claimed in claim 6, wherein the second sub-bands are [0Hz, 1100Hz], [1100Hz, 2500Hz], [2500Hz, 5500Hz] and [5500Hz, 11000Hz].

8. The method as claimed in claim 6, wherein the second sub-bands are determined by user input.

9. The method as claimed in claim 1 further comprising the step of filtering the sequences before application of auto-correlation, wherein only the components having amplitudes larger than a predetermined value are left unchanged while the others are set to zero.

10. The method as claimed in claim 1, wherein the audio stream is re-sampled by the steps of dividing the audio stream into chunks and joining two adjacent chunks into one block so that the blocks have samples overlapping with each other.

11. The method as claimed in claim 10, wherein the number of the samples in one chunk is 256.

12. The method as claimed in claim 1, wherein the energy summation of the n^{th} block within the i^{th} sub-band is derived from the following equation:

$$A_i(n) = \sqrt{\sum_{k=L_i}^{H_i} a(n,k)}\ ,$$

where L_i and H_i are lower and upper bounds of the i^{th} sub-band, and a(n,k) is an energy value (amplitude) of the n^{th} block at a frequency k.

13. The method as claimed in claim 1, wherein the first sub-bands are [0Hz, 125Hz], [125Hz, 250Hz] and [250Hz, 500Hz].

14. The method as claimed in claim 1, wherein the first sub-bands are determined by user input.

15. The method as claimed in claim 1 further comprising the step of determining beat onsets of the music soundtrack using the final estimated tempo.

16. The method as claimed in claim 15, wherein the beat onsets are determined by the steps of:

a) identifying a maximum peak in the sequence of the sub-band whose estimated tempo is the final estimated tempo;

b) deleting neighbors of the maximum peak within a range of the final estimated tempo;

c) identifying a next maximum peak in the sequence; and

d) repeating the steps b) and c) until no more peak is identified;

wherein all the identified peaks are the beat onsets.